

2023年统计建模大赛代码说明README文件

本项目目录结构

```

|
| README.md
|
├── Data
|   ├── GDP数据清洗.R
|   ├── 乡村恩格尔系数.R
|   ├── 发电量数据清洗.R
|   ├── 商品价格指数数据清洗.R
|   ├── 城镇化率数据清洗.R
|   ├── 城镇恩格尔系数数据清洗.R
|   ├── 就业人口数据清洗.R
|   ├── 社会消费品零售总额数据清洗.R
|   ├── 线性回归.R
|   ├── 财政支出数据清洗.R
|   └── 资本存量数据清洗.R
|
|   └── data
|       ├── cleanGDP78-17.csv
|       ├── cleanRPI78-17.csv
|       ├── cleanTRSCG78-17.csv
|       ├── clean乡村恩格尔系数78-17.csv
|       ├── clean产业结构指数78-17.csv
|       ├── clean发电量78-17.csv
|       ├── clean发电量78-17.xls
|       ├── clean城镇化率78-17.csv
|       ├── clean城镇恩格尔系数78-17.csv
|       ├── clean就业人口78-17.csv
|       ├── clean财政支出78-17.csv
|       ├── clean资本存量78-17.csv
|       ├── IPCAcoef.csv
|       ├── linearerror.csv
|       ├── linearresidual.csv
|       ├── result_prediction_label.mat
|       ├── yrprvcdata_sliced_all_log.mat
|       └── yrprvcdata_z_sliced.mat
|
├── DNN
|   ├── draw.py
|   ├── main.ipynb
|   ├── main.py
|   ├── requirements.txt
|   ├── result_prediction_label.mat
|   ├── yrprvcdata_sliced_all_log.mat
|   └── yrprvcdata_z_sliced.mat
|
├── images
|   └── results
|
```

```
第1个工具变量商品价格指数-GDP.jpg
第1个工具变量商品价格指数-theta.jpg
第2个工具变量社会消费品零售总额-GDP.jpg
第2个工具变量社会消费品零售总额-theta.jpg
第3个工具变量乡村恩格尔系数-GDP.jpg
第3个工具变量乡村恩格尔系数-theta.jpg
第4个工具变量产业结构指数-GDP.jpg
第4个工具变量产业结构指数-theta.jpg
第5个工具变量发电量-GDP.jpg
第5个工具变量发电量-theta.jpg
第6个工具变量城镇化率-GDP.jpg
第6个工具变量城镇化率-theta.jpg
第7个工具变量城镇恩格尔系数-GDP.jpg
第7个工具变量城镇恩格尔系数-theta.jpg
第8个工具变量财政支出-GDP.jpg
第8个工具变量财政支出-theta.jpg

—logs
  model_DNN_1000_2023-05-25-21-48-35.pth
  model_state_dict_DNN_1000_2023-05-25-21-48-35.pth

—models
  | DNN.py
  |
  |__pycache__
    DNN.cpython-37.pyc

—utils
  | Function.py
  | Load_data.py
  | Utils.py
  |
  |__pycache__
    Function.cpython-37.pyc
    Load_data.cpython-37.pyc
    Utils.cpython-37.pyc

—Images
  GDP.jpg
  IPCA+线性回归.jpg
  IPCA+线性回归.vsdX
  IPCA.jpg
  IPCA.vsdX
  Theta.jpg
  多层感知机.jpg
  我们的模型.jpg
  残差.eps
  残差.png
  神经元模型.jpg
  线性回归.jpg
  线性回归.vsdX

—I PCA
  | cleanGDP78-17.csv
```

```
cleanRPI78-17.csv
cleanTRSCG78-17.csv
clean乡村恩格尔系数78-17.csv
clean产业结构指数78-17.csv
clean发电量78-17.csv
clean发电量78-17.xls
clean城镇化率78-17.csv
clean城镇恩格尔系数78-17.csv
clean就业人口78-17.csv
clean财政支出78-17.csv
clean资本存量78-17.csv
IPCAcoef.csv
IPCAmain.m
IPCA_Gamma.m
linearerror.csv
linearresidual.csv
result_prediction_label.mat
yrprvcdata_sliced_all_log.mat
yrprvcdata_z_sliced.mat
```

└─Figures

```
residualanderror_Log.eps
residualanderror_Log.eps.png
```

原数据及其清洗

数据清洗参看Data文件夹下的R语言文件 原数据参看Data/data下的csv文件

IPCA

1. IPCA采用Matlab实现，目录下的**IPCAmain.m**即为主函数
 2. 依赖函数有**IPCA_Gamma.m**函数
 3. 使用数据为3个mat文件
-

DNN

1. DNN模型采用Pytorch框架实现，具体参看DNN文件夹下的文件目录及其结构
 2. 依赖包位于requirements.txt中
 3. 使用数据为3个mat文件
-

Images

论文中出现的图表
